Digital Communications Engineering 1

(COMM2108)

Source Formatting

Source Formatting

- Source formatting is concerned with how to achieve an accurate and efficient representation of the information to be communicated.
- Source information can have many different forms, all of which must be converted into a digital format in order to be transmitted over a digital communications system.

Source Formatting

- For example, the digital data from a PC/server is already in a digital format. However, the *efficiency* of the format may be improved using *compression*. This allows for a more compact representation which is useful for transmission or storage.
- In the case of text information, for example generated by a computer keyboard, there is a need to represent the letters A-Z in upper case and lower case, as well as the numbers 0-9 and punctuation marks. The solution is to use a digital *code*.

Source Formatting

- In order to encode text information, it is necessary to assign a unique binary code word to each letter, number, and punctuation mark.
- The best known and most widely used code for binary text encoding is the *ASCII Code* which uses a 7-bit code word.
- The ASCII code allows for up to $2^7 = 128$ unique binary code words.
- For example, the capital letter $A = 1000001$.

Source Formatting

- In the case of voice and video information the process of digital encoding is somewhat more complicated owing to the inherently *analogue* nature of these sources.
- For voice and video signals, the requirement is to convert an analogue waveform into an *accurate* and *efficient* digital representation of the signal.

Source Formatting

- A possible solution here is to use *Pulse Code Modulation (PCM)* which is a well established technique first developed in 1937 by a British engineer *Alec Reeves*.
- *PCM* is the basis of the modern digital revolution as it allows *any* analogue signal to be accurately and efficiently represented in a digital format.

Source Formatting

- Pulse Code Modulation (PCM) is actually a 3 step process comprising:
  - Sampling
  - Quantisation
  - Encoding
- *Sampling* involves taking a series of regularly spaced samples of the value of the input analogue waveform.
- This raises the question of how often should we sample and are there any restrictions on the sampling rate?

Source Formatting

- The answers are provided by the *Nyquist Sampling Theorem (1928)* which states that:

Any continuous bandwidth-limited signal of finite energy can be completely reconstructed (without any loss of information) from a series of regularly spaced samples T_s seconds apart provided that

$$\frac{1}{T_s} \geq 2 f_{\max}$$

where f_{max} is the maximum frequency component of the analogue signal.

Source Formatting

- It is worth noting that provided that this inequality is satisfied, there will be <u>no loss</u> of information, despite the “gaps” between the samples. This can be seen from the spectrum of the sampled signal.
- The original analogue signal can be recovered from the sampled signal by using a Low Pass Filter (LPF) of bandwidth f_{max}.
- In practice, an electronic circuit called a *sample-and-hold circuit* is used to sample the signal.

Source Formatting

- The output from the sample-and-hold circuit is a discrete-time signal consisting of a series of pulses whose amplitudes are given by the amplitude of the input analogue signal at the sampling instant.
- The sampled waveform is known as a *Pulse Amplitude Modulation (PAM)* signal.

Source Formatting

- A question arises as to what is the consequence of failing to satisfy the sampling theorem inequality, i.e. what happens if

$$\frac{1}{T_s} < 2 f_{\max}$$

- This situation is known as *undersampling* and leads to distortion, known as *aliasing*, being introduced in to the sampled signal.
- Undersampling leads to an overlap of the frequency components in the sampled spectrum known as *spectral folding*.

Source Formatting

- In reality, no signal is completely bandwidth limited, so there is a need to ensure that the input analogue signal has no frequency components above f_{max}.
- It is necessary to use a filter before the sampling in order to explicitly remove any frequency components above f_{max}.
- This filter is known as an *anti-aliasing filter* as it helps prevent aliasing due to undersampling.

Source Formatting

- Furthermore, since filters with an infinitely sharp cut-off are not physically realisable, there is an additional requirement to *oversample*.
- This leads to an "engineering" version of the sampling theorem which states that the sampling frequency f_s is given by

$$\begin{aligned} f_s &= \frac{1}{T_s} \\ &\geq 2.2\, f_{\max} \end{aligned}$$

- In practice, one would sample between 2 to 3 times f_{max}.

Source Formatting

- *Question*: Is the output from the sample-and-hold circuit a *digital* signal?
- *Answer*: No, most definitely not!!
- The reason is that the samples values are continuous since they are given by the amplitude of the input analogue waveform at the sampling instant.
- Consequently, these analogue sample values are incompatible with a digital system since in a digital system only a *finite* number of values are allowed.

Source Formatting

- Another step known as *quantisation* is required whereby the continuously variable (i.e. analogue) sample values are converted into a <u>finite</u> number of permitted values.
- *Quantisation* is the process that segregates the sample values into ranges and assigns a unique identifier to each range.
- Whenever a sample value falls within a range, the output is the discrete identifier assigned to the range.

Source Formatting

- The final step in the PCM process is binary *encoding* where a unique binary code word is assigned to each of the sample ranges.
- If a *n*-bit binary code word is used, the maximum number of quantisation levels (and hence the maximum number of permissible sample values) that can be represented is simply 2^n levels.

Source Formatting

- Example: In telephony grade (or toll-quality) voice PCM the human voice is considered to have a frequency spectrum between *300* and *3400 Hz.*
- Firstly, the voice waveform is *bandwidth limited* by passing it through an anti-aliasing filter with a cut-off frequency of *3400 Hz*.
- The bandwidth limited voice waveform is then *oversampled* at rate of *8000* samples per second.
- The sampled voice waveform is then quantised using a $2^8 = 256$ *level* quantisation scheme.

Source Formatting

- An 8-bit binary code word is used to represent the quantised voice sample.
- The resulting PCM bit stream therefore has a *bit rate* of:

8000 samples/s x 8 bits/sample = 64,000 bits per sec

= 64 kbps

- This bit rate represents a fundamental unit in the design of digital telephone systems as it represents the bit rate required to support one PCM encoded voice channel.

Source Formatting

- Quantisation is an <u>irreversible</u> process. In other words, the rounding process that took place at the transmitted cannot be "undone" at the receiver.
- Consequently, the analogue waveform reconstructed at the receiver will <u>never</u> be an exact replica of the original input analogue waveform.
- This distortion introduced into the reconstructed waveform is considered to be a form of noise, known as *quantisation noise*.

Source Formatting

- The power (or the rms value) of this quantisation noise associated with the PCM process can be calculated and measured.
- This leads to the concept of a signal-to-quantisation noise ratio (*SN_qR*) where

$$SN_qR = \frac{Signal\ Power}{Quantisation\ Noise\ Power}$$

$$= \frac{\overline{S}}{\overline{N_q}}$$

Source Formatting

- The mean signal power can be shown to be given by

$$\overline{S} = \frac{M^2 - 1}{12} q^2$$

- where *q* is the quantisation interval and *M* is the number of quantisation levels.
- The mean quantisation noise power can be shown to be given by

$$\overline{N_q} = \frac{q^2}{12}$$

Source Formatting

- Consequently, the signal-to-quantisation noise power ratio is given by

$$SN_qR = M^2 - 1$$

- For a large values of *M*, the following approximation

$$SN_qR \approx M^2$$

is often used.

Source Formatting

- It is worth noting that

$$SN_qR = M^2 - 1$$

- represents the <u>average</u> signal-to-quantisation noise power ratio.
- The <u>peak</u> signal-to-quantisation noise power ratio is

$$(SN_qR)_{peak} = 3M^2$$

Source Formatting

- Expressing the SN_qR in terms of dB gives

$$SN_qR = 20\log_{10}(M) \qquad (dB)$$

$$(SN_qR)_{peak} = 4.8 + SN_qR \qquad (dB)$$

- If the M quantisation levels are encoded into n-bit words where $n = log_2M$ then

$$\begin{aligned}(SN_qR)_{peak} &= 3M^2 \\ &= 3(2^n)^2\end{aligned}$$

Source Formatting

- Expressed in *dB* gives

$$(SN_qR)_{peak} = 4.8 + 6n \qquad (dB)$$

- This expression is known as the "*6 dB*" rule which illustrates the significant performance characteristic of PCM where there is an additional *6-dB* improvement in SN_qR obtained for each bit added to the PCM word.
- The SN_qR is very much dependent on the ratio of peak to mean signal power.

Source Formatting

- If we define the ratio of peak S_{peak} to mean S_{mean} signal power as

$$\alpha = \frac{S_{peak}}{S_{mean}}$$

- Then

$$SN_qR = \frac{3(2^{2n})}{\alpha}$$

- Expressed in *dB* gives

$$SN_qR = 4.8 + 6n - \alpha_{dB} \qquad (dB)$$

Source Formatting

- For a sinusoidal signal α = *2 (i.e. 3dB)* while for human speech α = *10 dB*.
- As a rule-of-thumb for a *n*-bit PCM voice system, the following expression is often used

$$SN_qR \approx 6(n-1) \qquad (dB)$$

Source Formatting

- Unfortunately, the SN_qR is <u>not</u> a constant, it varies with the level of signal power (see handout which shows the SN_qR experienced by a sinusoidal signal quantised with 256 equally spaced levels).
- For a full load signal, that sweeps through all 256 levels, the $SN_qR = 50\ dB$, while an input signal down 20dB on full load has a $SN_qR = 30\ dB$.
- However, as the dynamic range of voice signals can be as large as 40 dB and given that "toll-quality" speech requires a SN_qR of at least *30 dB*, the performance of a linear quantiser with 256 levels can only handle part of the dynamic range for human speech.

Source Formatting

- If the number of bits *n* were increased from *8* to *12* bits, then an input signal level of *–40dB* (relative to full load) will experience a $SN_qR = 34\ dB$.

- However, $2^{12} = 4096$ quantisation levels would be required and until recently it was not possible to achieve such precision in ADC circuits.

- Therefore a technique called *companding* was developed to improve the performance of *8*-bit quantising.

Source Formatting

- *Companding* is equivalent to *nonuniform* quantising where a variable step size is used.
- In *nonuniform* quantising the step size between adjacent quantisation levels is deliberately tapered so that small step sizes are used for low level signals and larger step sizes for higher level signals.
- *Companding* achieves the same result by *compressing* the signal (using a nonlinear amplitude characteristic) prior to linear quantisation and then *expanding* the reconstructed signal (with the inverse amplitude characteristic).
- In both cases, the strategy is to attempt to maintain a constant SN_qR for all signal levels.

Source Formatting

- If uniform quantisation is used, the signal should be compressed such that increasing the input signal by a given factor increases the output signal by a corresponding additional constant.
- This implies a *logarithmic* relationship for the compression characteristic, i.e.

$$y = \log x$$

- Since signals can take on negative as well as positive values, the logarithmic characteristic must be reflected to form an *odd symmetric* function.

Source Formatting

- Furthermore, the characteristic must be continuous across zero volts.
- Therefore the compression characteristic has the form of two logarithmic function joined by a linear section.
- There are two compression characteristics that been standardised for use by the world's telephones companies, these are the *μ-law* and the *A-law* characteristics.

Source Formatting

- The μ*-law* compression standard is used in North America and Japan and has the following form:

$$y = \operatorname{sgn}(x)\frac{\ln\left(1+\mu|x|\right)}{\ln(1+\mu)} \qquad 0 \le |x| \le 1$$

- Originally μ = *100* for use with a *7*-bit converter, however it was later changed to μ = *255* for an *8*-bit converter. Note that μ = *0* corresponds to linear quantisation.

Source Formatting

- The *A-law* compression standard is used in Europe and the rest of the world and has the following form:

$$y = \operatorname{sgn}(x)\frac{1+\ln(A|x|)}{1+\ln(A)} \qquad \tfrac{1}{A} \le |x| < 1$$

$$= \operatorname{sgn}(x)\frac{A|x|}{1+\ln(A)} \qquad 0 < |x| \le \tfrac{1}{A}$$

Source Formatting

- The standard value of $A = 87.6$ gives a $24\ dB$ improvement in SN_qR over linear PCM for small signals (i.e. $|x| < 1/A$) and a constant $SN_qR = 38\ dB$ for large signals (i.e. $|x| > 1/A$).

- The dynamic range for large signals is *39 dB* which is equivalent to an *11*-bit quantisation scheme.

- The *A-law* compression characteristic is normally implemented as a 13-segment piece-wise linear approximation.

Source Formatting

- In practice 16 segments are used but with 4 segments co-linear near the origin.
- For 8-bit PCM, one bit is used for polarity, 3 bits indicate which segment the sample lies on and 4 bits provide the location on the segment.
- The μ*-law* (with $\mu = 255$) gives slightly improved SN_qR for voice signals but it has a smaller dynamic range.
- The μ*-law* compression characteristic is implemented as a 15-segment piece-wise linear curve.

Source Formatting

- *A-law* and *μ-law* companded PCM have been adopted by the *ITU-T* as international toll-quality telephony standards (Rec. G.711).

- The G.711 PCM communication has a Mean Opinion Score (MOS) speech quality measured by subjective testing of 4.3 on a scale of 0 to 5.

Source Formatting

- If the PCM code words are received and decoded <u>without</u> error, then the *SNR* of the decoded signal is essentially given by SN_qR.
- In the presence of channel and/or receiver noise, bit errors may occur. The effect that a bit error has on the *SNR* of the decoded signal depends on which bit is detected in error.
- An error in the least significant bit (LSB) in the PCM code word will introduce an error in the decoded signal equal to one quantisation level.
- An error in the most significant bit (MSB) will introduce an error of many quantisation levels.

Source Formatting

- Before deriving an expression for the *SNR* of the decoded PCM signal it is necessary to state a few assumptions:
  - Assume that the probability of more than one bit error occurring in a single *n*-bit PCM code word is negligible.
  - Assume that the probability of error P_e is the same for all bits in the PCM code word.

Source Formatting

- The possible errors in the decoded PCM signal are:

$$\varepsilon_1 = q$$

$$\varepsilon_2 = 2q$$

$$\varepsilon_3 = 4q$$

$$\vdots$$

$$\varepsilon_n = 2^{n-1}q$$

- where the subscripts *1, 2, …, n* denotes the position or significance of the PCM code word bit.

Source Formatting

- The mean square decoding error is given by mean square of all the possible errors multiplied by the probability of the error occurring in the code word

$$\overline{\varepsilon_{de}}^{2} = P_e\left[(q)^2 + (2q)^2 + ... + (2^{n-1}q)^2\right]$$
$$= P_e(q)^2\left[4^0 + 4^1 + 4^2 + ... + 4^{(n-1)}\right]$$

- Using the following result for geometric series

$$S_n = a + ar + ar^2 + ... + ar^{n-1}$$
$$= \frac{a(r^n - 1)}{r - 1}$$

Source Formatting

- The mean square decoding error is

$$\overline{\varepsilon_{de}}^{2} = \frac{P_e q^2 (4^n - 1)}{3}$$

- Since the errors due to the quantisation process and the transmission errors are statistically independent, their power or mean square values can be summed to give

$$SNR = \frac{\overline{S}}{\overline{N_q} + \overline{\varepsilon_{de}}^{2}}$$

Source Formatting

- Using the previous results for the mean signal power and mean quantisation noise power

$$\overline{S} = \frac{M^2 - 1}{12} q^2$$

$$\overline{N_q} = \frac{q^2}{12}$$

- This gives

$$SNR = \frac{M^2 - 1}{1 + 4(M^2 - 1)P_e}$$

Source Formatting

- In terms of <u>average</u> signal power

$$SNR \approx \frac{M^2}{1+4(M^2-1)P_e}$$

- In terms of <u>peak</u> signal power

$$SNR \approx \frac{3M^2}{1+4(M^2-1)P_e}$$